Martins et al. Forest Ecosystems (2016) 3:17
DOI 10.1186/s40663-016-0073-8

Forest Ecosystems

RESEARCH Open Access

A multi-scale modelling framework to guide management of plant invasions in a transboundary context

João Martins[1*], David M. Richardson[2], Renato Henriques[3], Elizabete Marchante[4], Hélia Marchante[4,5], Paulo Alves[1,6], Mirijam Gaertner[2,7], João P. Honrado[1,6] and Joana R. Vicente[1,6]

Abstract

Background: Attention has recently been drawn to the issue of transboundary invasions, where species introduced and naturalized in one country cross international borders and become problematic in neighbouring countries. Robust modelling frameworks, able to identify the environmental drivers of invasion and forecast the current and future potential distribution of invasive species, are needed to study and manage invasions. Limitations due to the lack of species distribution and environmental data, or assumptions of modelling tools, often constrain the reliability of model predictions.

Methods: We present a multiscale spatial modelling framework for transboundary invasions, incorporating robust modelling frameworks (Multimodel Inference and Ensemble Modelling) to overcome some of the limitations. The framework is illustrated using *Hakea sericea* Schrad. (Proteaceae), a shrub or small tree native to Australia and invasive in several regions of the world, including the Iberian Peninsula. Two study scales were considered: regional scale (western Iberia, including mainland Portugal and Galicia) and local scale (northwest Portugal). At the regional scale, the relative importance of environmental predictors sets was evaluated and ranked to determine the main general drivers for the species distribution, while the importance of each environmental predictor was assessed at the local scale. The potential distribution of *H. sericea* was spatially projected for both scale areas.

Results: Model projections for western Iberia suggest that a large area is environmentally suitable in both Portugal and Spain. Climate and landscape composition sets were the most important determinants of this regional distribution of the species. Conversely, a geological predictor (schist lithology) was more important in explaining its local-scale distribution.

Conclusions: After being introduced to Portugal, *H. sericea* has become a transboundary invader by expanding in parts of Galicia (Spain). The fact that a larger area is predicted as environmentally suitable in Spain raises concerns regarding its potential continued expansion. This highlights the importance of transboundary cooperation in the early management of invasions. By reliably identifying drivers and providing spatial projections of invasion at multiple scales, this framework provides insights for the study and management of biological invasions, including the assessment of transboundary invasion risk.

Keywords: Drivers of invasion, *Hakea sericea*, Multimodel inference, Transboundary invasion management, Species distribution models

* Correspondence: joaofilipepm@gmail.com
[1]InBIO / CIBIO - Centro de Investigação em Biodiversidade e Recursos Genéticos, Universidade do Porto - Campus Agrário de Vairão, Rua Padre Armando Quintas, n° 7, 4485-661 Vairão, Portugal
Full list of author information is available at the end of the article

Background

Invasion by alien species is a major threat to ecosystems worldwide. Many invasive species cause substantial changes to ecosystem structure and functioning (Vilà et al. 2011), provision of ecosystem services (Vilà et al. 2010) and the broader economy (Pimentel et al. 2005; Vilà et al. 2010). Some invasive species potentially promote irreversible regime shifts (Gaertner et al. 2014), and cause biotic homogenization synergistically with other drivers of global change (Thuiller 2007). Therefore, biological invasions as a global change process, and a worldwide problem, must be managed through international cooperation.

The concept of transboundary pollution, where pollution originating in a given country negatively impacts the environment in another country, has long been a part of environmental sciences, with a history of application and evaluation of practical measures which show that international cooperation is key in dealing with this problem (United Nations 1997; EEA 1999, EEA 2015). The parallel concept of transboundary invasions, where a species introduced in a given country expands into neighbouring countries has so far received much less attention, despite being discussed in recent publications (e.g. Hulme 2015; Roques et al. 2016) and legislation calling for concerted action between countries (European Parliament and Council of the European Union 2014). Therefore, progress in this area demands that traditional risk assessment methodologies be expanded to include multiple countries (Hulme 2015). In this paper we apply species distribution models to predict invasion patterns in a transboundary context.

Despite the conceptual similarities to other international environmental issues and concerns, biological invasions are a very special case, and several particularities of the phenomenon need to be taken into consideration before addressing it. The invasion process involves several stages (transport, introduction, establishment, and spread), with sequential barriers (e.g. geographical, survival, reproductive, dispersion) that must be overcome by a species in order to reach, survive and spread in a new non-native territory (Richardson et al. 2000; Blackburn et al. 2011). The success of invasive species in a new territory depends on their invasiveness (i.e. features of the organism that define its ability to invade), and on the invasibility of the local ecological systems (i.e. characteristics of those systems that determine the susceptibility to invasion; Richardson et al. 2011). Invasion patterns and processes exhibit spatial dependence, and the relative importance of different sets of environmental factors (e.g. climate, landscape composition and structure, disturbances) to explain species distributions varies across spatial scales (Rouget and Richardson 2003; Pearson et al. 2004; Guisan and Thuiller 2005; Vicente et al. 2014). Consequently, attempts to understand and predict invasion processes must always consider the spatial scale(s) of the study system (Pauchard and Shea 2006; Theoharides and Dukes 2007), and produce more informative predictions of invasion than traditional models (Vicente et al. 2011).

Species distribution models (SDMs) statistically relate the distribution of a given species with environmental factors, improving the understanding and prediction of the potential distribution of species in a specific territory (Elith and Leathwick 2009). SDMs have been widely used in the field of biological invasions (e.g. Peterson 2003; Broennimann et al. 2007; Vicente et al. 2010, 2011; Fernandes et al. 2014). Despite their usefulness, a recurrent difficulty in applying SDMs is the likelihood of model overfitting due to a high ratio between the number of environmental predictors and the often-scarce species occurrence data (Guisan and Thuiller 2005). To overcome this problem, information-theoretic approaches such as Multimodel Inference (Burnham and Anderson 2002) can be applied. Multimodel Inference allows the comparison of different hypotheses on invasion predictors, by comparing and ranking a series of competing models that reflect different hypotheses for explaining the species' distribution, and by measuring each model' contribution to explain the observed data. Additionally, weighting each model by its importance to explain the original dataset (i.e. model averaging) allows us to obtain a consensus spatial projection (Vicente et al. 2010). Overall, this procedure paves the way for more accurate models and projections while testing the effect of a larger set of environmental predictors.

Another important difficulty when applying species distribution models arises as different modelling techniques can produce very different outputs, even when models are calibrated with the same occurrence and environmental data (Pearson et al. 2006). A panoply of modelling techniques is currently available (Araújo and New 2007; e.g. Guisan et al. 2002; Olden et al. 2008), confounding the choice of the most appropriate technique to reach a given goal. To avoid variability in predictions, ensemble-forecasting modelling can be applied, by fitting a series of models using multiple techniques and then combining the predictions into a consensus prediction (weighted by the accuracy of the different methods; Araújo and New 2007).

The shortage of data on the distribution of invasive species can to some extent be overcome by supplementing data from research and monitoring programs with data acquired from citizen science initiatives (Crall et al. 2010). However, data from citizen science programs, as exemplified by the web mapping platform available at www.invasoras.pt (invasoras.pt 2014), are often collected only in the form of presence records, making it

necessary to generate pseudo-absence records, following the most consensual procedures to avoid biasing the accuracy of model predictions (Wisz and Guisan 2009).

We propose a modelling framework to identify and rank multiscale environmental predictors of transboundary distribution of invasive species based on scarce occurrence data (a common limitation in biodiversity modelling studies; Lomba et al. 2010; Vicente et al. 2011). This allows the use of data from a more invaded country (Portugal in our example) to predict areas of potential risk in neighbouring countries with slight or no current invasion (Spain in our example), provided that model transferability principles are observed (Elith and Leathwick 2009). In the proposed framework, the first step is to identify areas of high invasion risk or of particular conservation importance at a regional scale (western Iberian Peninsula in our study case). The second step involves zooming in to the areas selected in step 1 (northwest Portugal in our study) in order to obtain more fine-grained predictions of potential invasion and to rank its driving factors. A key feature of the approach is that the different scales of analysis are connected in that equivalent variables are used, each analysis does not try to be a scaled version of the other, allowing the use of more suited techniques and predictors for each scale. We illustrated the framework for the transboundary invasion by *Hakea sericea* Schrad., an invasive shrub invading mainland Portugal and currently spreading to neighbouring areas in Spain. We obtained spatial projections of potential distribution for both scales, and implications for transboundary management of invasive species are discussed.

Methods

Study areas

The proposed framework was applied in two nested geographic areas (Fig. 1): a regional-scale area including Portugal and Galicia, covering ca. 133 000 km^2, and a local area in the Minho region, Northwest Portugal, covering ca. 2972 km^2. The use of nested areas, with different grain and extent, allows for the identification of the main factors underlying species distributions at multiple spatial scales (e.g. see Lomba et al. 2010; Vicente et al. 2011, 2014).

The regional-scale area, located in the western part of the Iberian Peninsula, includes the transition between the Eurosiberian and the Mediterranean biogeographic regions of Europe, with Galicia and northwest Portugal representing the Eurosiberian areas (Rivas-Martínez et al. 2004). The large variability in topography, geology, soils and land cover, along with the transitional biogeographic situation, results in a highly heterogeneous area in terms of environmental and socio-ecological contexts.

The Minho region, in northwest Portugal, is a particularly well-studied and data-rich area for both invasive species occurrence and environmental data (e.g. Vicente et al. 2010, 2011; Fernandes et al. 2014). It is located within the area predicted as suitable for the study species by the regional scale model (see below), hence its choice for the fine-scale component of this work. This area has high climatic heterogeneity, with gradients that constrain not only native biodiversity (e.g. Lomba et al. 2010) but also invasive plant species richness (Vicente et al. 2010) and the potential distribution of individual invaders, with climate acting as the primary driver (Vicente et al. 2011). To allow for the identification of potentially important non-climatic predictors, a climatically homogeneous area (Fig. 2c) was selected within the region, as performed in Vicente et al. (2010); see "Model Calibration" section.

Study species and occurrence data

Hakea sericea Schrad. (silky hakea; Proteaceae) is a woody shrub or small tree that is native to south-eastern Australia where it occurs in areas of dry sclerophyll forest and heathlands on coastal regions and adjacent ranges, from south-eastern Queensland to south-eastern New South Wales (Barker et al. 1999). This fire-adapted species is invasive in New Zealand, mainland Portugal and islands (Madeira), and South Africa (Rejmánek and Richardson 2013). The invasion dynamics of the species have been intensively studied in South African fynbos vegetation (e.g. van Wilgen and Richardson 1985; Richardson et al. 1987; Le Maitre et al. 2008). Invasion by *H. sericea* in fynbos causes major changes to the fuel characteristics of this fire-prone shrubland vegetation (van Wilgen and Richardson 1985). Due to the capacity to sustain and promote fire, invasion by *H. sericea* can lead to severe ecological and socioeconomic regime shifts (Gaertner et al. 2014). Although such major impacts have yet to be quantified outside South African fynbos, *H. sericea* has the potential to become more widespread and abundant in Southern Europe and other areas with Mediterranean-type climate, as suggested by its aggressive invasiveness in South Africa.

H. sericea is a serotinous species, whose seeds are retained in woody fruits. Once mature, the fruit's thick woody walls offer protection to the seeds against heat (Brown and Whelan 1999) and granivores (Groom and Lamont 1997). This, together with a slow decrease in germination rates over time (Richardson et al. 1987; Brown and Whelan 1999), results in a large canopy-stored seed bank formed by seeds produced in a given year and viable seeds from previous years, which are released only after the death of the plant

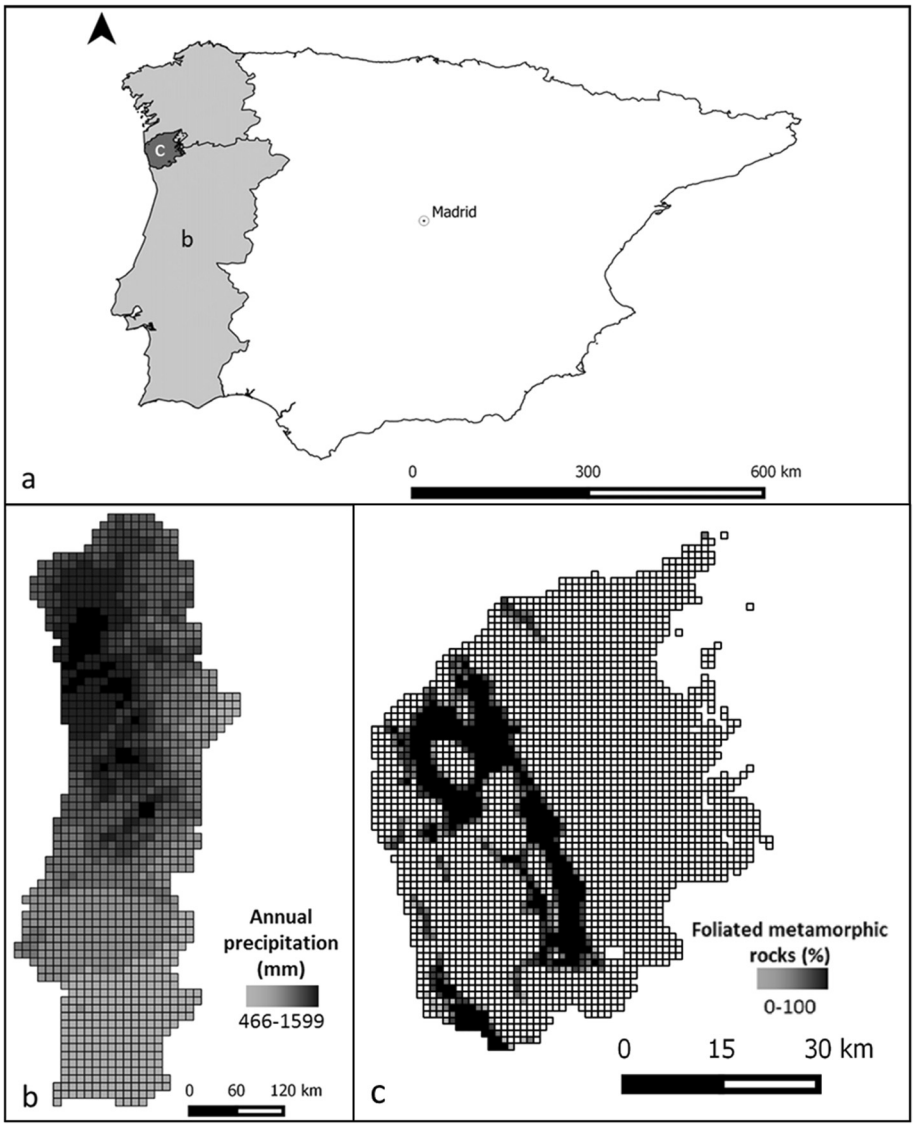

Fig. 1 Study areas for the determination of the drivers of distribution of *Hakea sericea* in the context of the Iberian Peninsula (**a**), regional-scale area with values of annual precipitation (**b**) and local-scale study area with values of the percentage area of foliated metamorphic rocks per cell (**c**). For the regional-scale area (light grey, **b**) the administrative divisions are presented as Portugal (southern subdivision) and Galicia (northern subdivision)

(Richardson et al. 1987). Therefore, an event such as a wildfire can result in the sudden release of a very large number of seeds.

Hakea sericea has been cultivated in Portugal as a hedge plant at least since the 1930s, and is known to have naturalized in natural vegetation in the 1940s (Espírito Santo and Arsénio 1999). It has become highly invasive in some areas (Marchante et al. 2014) and has spread to at least one location in northwest Spain (Pulgar Sañudo 2006), thereby becoming a transboundary invasion. *Hakea sericea* is listed in Portuguese legislation as an invasive species (Ministério do Ambiente 1999), and is considered a potential invader in Spain (Ministerio de Agricultura Alimentacion y Medio Ambiente 2011). Brunel et al. (2010) consider *H. sericea* as an emerging invasive alien plant in the Mediterranean Basin, with a potentially severe impact on the environment.

Occurrence records for *H. sericea* were obtained from previous field surveys, from published studies (Vicente et al. 2010), and from the citizen science web platform invasoras.pt (2014). A total of 53 presence records for the regional scale (10 km × 10 km grid cells) and 108 records for the local scale (37 presences and 71 absences; 1 km × 1 km grid cells) were obtained and used to calibrate the models.

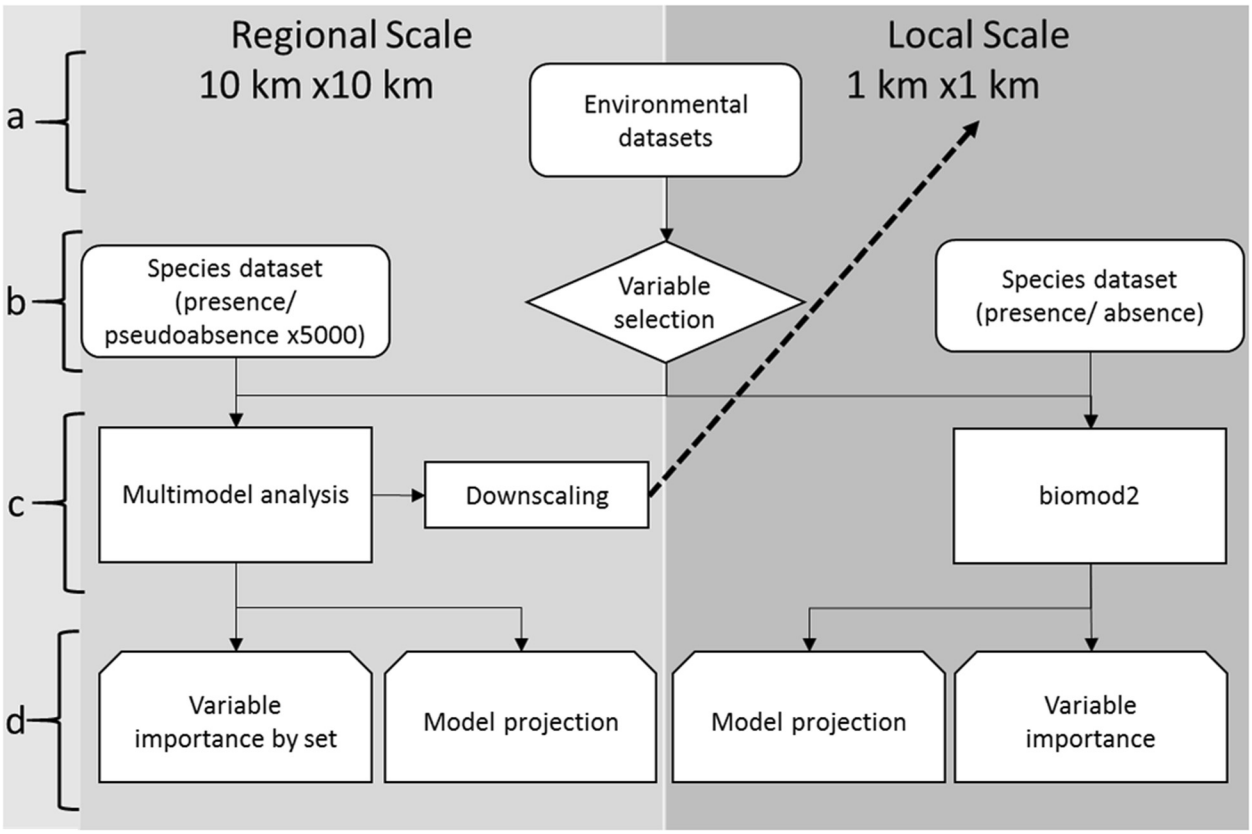

Fig. 2 Multi-scale modelling framework for the analysis of drivers and patterns of *Hakea sericea* distribution in the Iberian Peninsula. The analysis was performed at two spatial scales (step **a**, regional and local). Multimodel inference analysis was applied at the regional scale analyses (**b**, **c**) to infer the importance of each set of predictors in determining the distribution of the species, for spatial projection (**d**), and in a downscaling exercise to establish a local-scale area (**c**). At the local scale an ensemble model was calibrated using the *biomod2* R package (**c**), and predictor importance scores were obtained (**d**). For both areas, spatial projections for the potential distribution of *H. sericea* were obtained (**d**)

Analytical framework

The proposed multi-scale modelling framework was applied to assess the transboundary invasion by *Hakea sericea* as described in Fig. 2.

Starting from an initial dataset of 65 environmental predictors (step a), a subset of 16 predictors was selected (step b) and classified into four groups that reflect different types of environmental factors (see below). The four groups of predictors were used to calibrate competing models in a multimodel inference analysis using *H. sericea* presence records and multiple sets of randomly selected pseudo-absences (Wisz and Guisan 2009; see below). The models were then used (step c) to determine which environmental sets were most important in explaining the distribution of *H. sericea* in the regional study area and to perform a spatial projection of the species' distribution (step d). This projection supported the selection of the local study area, by applying a downscaling procedure for the local scale area using the regional scale models (step c). Since a larger number of occurrence records were available for the local area (including confirmed absences), it was possible to apply ensemble modelling using the *biomod2* package (Thuiller et al. 2009, 2015), implemented in the R software (R Development Core Team 2014; step c). The ensemble modelling was also used to obtain an importance score for each environmental predictor, and to project the potential spatial distribution of the species for the local study area (step d).

Statistical procedures throughout the whole workflow were performed using R 3.1.0 (R Development Core Team 2014). Spatial operations were performed using ArcMap 10.2 (ESRI 2014) and QGIS 2.2 (QGIS Development Team 2014). Further methodological details on model development are provided in the following sections.

Model calibration, evaluation and spatial projection

Predictor selection

For the regional scale, we started with 65 environmental predictors reflecting four main types of environmental conditions: climate, landscape composition, landscape

structure and lithology. Based on ecological expert knowledge of the species and the study area, and applying a pairwise Spearman correlation analysis (to avoid multicollinearity), a final dataset of 16 (Spearman correlation ≤ 0.7) environmental predictors (four per type) was obtained (Table 1). To allow comparability across scales, a dataset with equivalent predictors was obtained for the local scale. At this scale, data on fire history was available, and so a predictor reflecting burnt area was added to the final dataset. The following set of variables was available: maximum burnt area of each cell in the last 10 years, average fire recurrence, and total number of fires, all of which presented correlation values (Spearman rho) above 0.9 among them. As such, the first variable was selected, as it provides information on both the burnt area and an indication of fire recurrence when the cell is completely burned more than once (by having a value over 100 %). The 17 predictor dataset was tested using a pairwise Spearman correlation analysis, and all predictors with Spearman correlation ≤ 0.7 were selected, retaining those predictors with the greatest ecological relevance for the species. A final dataset of 13 predictors was obtained for the local-scale area, including at least two predictors from each set from the regional-scale, and the fire predictor (Table 1).

Regional scale model

Multimodel Inference was applied to determine the relative importance of a set of predictors (see Appendix II for the list of predictors used in each model) on the distribution of the test species, by calibrating Generalized Linear Models (GLMs, calibrated with Poisson variance and log link function), using the Akaike Information Criterion (AIC) to assess how much each model was supported by the initial occurrence data. An adaptation of AIC for small sample sizes (AIC_c) was used; for each candidate model, the AIC_c difference was calculated ($\Delta i = AIC_{c\ \text{initial}} - AIC_{c\ \text{minimum}}$), allowing the comparison among all competing models. Finally, the Δi values were used to derive Akaike weights (w_i), representing the likelihood that a given model is the best approximating model, given the model and data sets. To assess explanatory accuracy, Nagelkerke's R^2 was calculated (Nagelkerke 1991). The w_i values were

Table 1 Predictors used in model calibration by predictor class, their description and ecological rationale for the selection

Class	Predictor	Description	Ecological rationale
Climate	MinTemp	Minimum temperature of the coldest month	Climate is expected to be the main factor in shaping species' distribution at large scales (Pearson et al. 2002), and previous studies have indicated climate predictors as the most important drivers of invasive species richness in the local-scale study area (Vicente et al. 2010).
	TempRan	**Temperature annual range**	
	AnnPrec	**Annual precipitation**	
	PrecSea	Precipitation seasonality (Coefficient of variation)	
Landscape composition	**pUrbanA**	**Urban areas cover (%)**	Land cover/use determine suitable habitat availability, thereby controlling alien invasion, and more invasive species find suitable conditions in man-made habitat (Song et al. 2005). Also, the greater the compositional diversity of a landscape, the more alien invasive species can find suitable conditions there (Pino et al. 2005).
	pAgrico	**Agriculture cover (%)**	
	pArtFor	**Artificial forests cover (%)**	
	pShrubs	Shrubland cover (%)	
Landscape structure	**NumPatc**	**Number of patches**	Spatial configuration and variability of the landscape affect species richness (Dufour et al. 2006), with landscape fragmentation increasing vulnerability to invasion (Le Maitre et al. 2004).
	mShaInd	**Mean Shape Index**	
	mPerAreR	**Mean Perimeter-to-Area Ratio**	
	shDiInd	**Shannon Diversity Index of Land Cover classes**	
Lithology	**IgnRock**	**Felsic and intermediate Igneous Rocks (%)**	The test species has adaptations related to nutrient absorption (Sousa et al. 2007). We expect that this might be reflected in differential suitability and invasive potential among different bedrocks, an expectation supported by empirical observations in some locations within our study area.
	MetRock	Foliated metamorphic rocks (%)	
	DetSedD	**Detritic sedimentary rocks and superficial sedimentary deposits (%)**	
	lithSDI	**Shannon Diversity Index of lithological categories**	
Fire	**pMaxBurn**	**Maximum burnt area per cell (%)**	*Hakea sericea* possesses fire adaptations (Groom and Lamont 1997; Brown and Whelan 1999), and invasion by this species potentially causes impacts on fire regimes (van Wilgen and Richardson 1985).

All predictors were used for the regional scale analysis, predictors in bold were selected for the local scale analysis (for more information see Appendix I)

used as weights to obtain an average model, for spatial projection of the outputs.

To calibrate the regional model, pseudo-absences were randomly selected in the study area. To avoid the model being skewed due to the pseudo-absence selection, a bootstrapping procedure, where the full presence set was used together with a random set of pseudo-absences (with the same size as the presence set), was applied to calibrate the GLMs. This process was repeated 5000 times (each time with a different pseudo-absence set). The AIC_c and related values were calculated for each repetition, and averaged over the 5000 repetitions. Furthermore, a null competing model (random model) was calibrated assuming that the considered hypotheses have no effect over the species' distribution.

Model evaluation was performed through the widely used Area Under the Receiver Operating Characteristic curve (ROC-AUC), as well as with the recently developed Boyce Index (Hirzel et al. 2006), implemented in the ecospat R package (Broennimann et al. 2015). The model predictions were finally converted into presence/absence using a ROC plot-based approach (threshold value corresponding to the point where the ROC curve is closest to the (0, 1) coordinates as discussed in Liu et al. (2005). ROC plot was performed using the pROC R package (Robin et al. 2011).

Finally, the projected average model was used for the selection of the local-scale area, by applying direct downscaling (e.g. Araújo et al. 2005; Fernandes et al. 2014). This was performed using models calibrated at the regional scale (10 km × 10 km grid) to perform a spatial projection at the local scale (1 km × 1 km grid), and using this projection to predict potential presence areas at the local scale.

Local scale model

A single class of model (GLMs) was applied at the regional scale, while ensemble models were used at the local scale, developed using the *biomod2* package (Thuiller et al. 2009, 2015; see Additional file 1 for details on ensemble modelling and its implementation) implemented in R. A total of 310 models for *H. sericea* were calibrated/fitted, using the 10 modelling techniques available in *biomod2*: GLM generalised additive models (GAM; Hastie and Tibshirani 1990), multivariate adaptive regression splines (MARS; Friedman 1991), classification tree analysis (CTA; Breiman et al. 1984), mixture discriminant analysis (MDA; Hastie et al. 1994), artificial neural networks (ANN; Ripley 1996), generalised boosted models (GBM; Ridgeway 1999), random forests (Breiman 2001), Surface Range Envelope (SRE; Busby 1991) and MaxEnt (Phillips et al. 2004).

Model evaluation was performed using the AUC metric using a cross-validation procedure (80 % of the data used for calibration/20 % for evaluation), with 30 repetitions.

The final model was obtained by an ensemble of the predictions of models with AUC above 0.7, using a weighting approach (Thuiller et al. 2015). A final evaluation of the ensemble model performance was based both on the AUC value and on the Boyce Index.

The importance of each predictor was estimated for the ensemble model prediction using the "variables_importance" function available in *biomod2* (Thuiller et al. 2015).

Results

Potential distribution of *Hakea sericea*

Figure 3 presents the potential distribution, based on the model results, for both the regional (a) and local (b) scales. The regional scale prediction was obtained by projecting the average model over the study area and converting this into a binary presence/absence prediction. At local scale the prediction was obtained from the projection of the ensemble of models from *biomod2*. Of a total of 1330 grid cells at the regional scale, 304 (22.9 %) were identified as having suitable conditions for *H. sericea* (Fig. 3a). These potential presences are located mostly in the western part of the study area, especially in the western half of central and northern Portugal (250 grid cells), extending northward to southwest Galicia (Spain, 54 grid cells – versus a single currently documented occurrence). Overall, the projection of the averaged regional model held very good predictive power (AUC = 0.882; Boyce Index = 0.713; Liu et al. 2005; Hirzel et al. 2006).

Only 180 cells (6.1 %) were predicted as potential presence of *H. sericea* in the local study area (Fig. 3b). The majority of these predicted presences are located in the northwest corner of the area, characterised by the presence of schist lithology (see Fig. 1c). The evaluation of model performance again indicates a high predictive power (AUC = 0.9; Boyce Index = 0.958; Liu et al. 2005; Hirzel et al. 2006).

Multi-scale drivers of the distribution of *Hakea sericea*

Table 2 indicates the importance of each set of predictors (for the regional scale) and each predictor (for the local scale). The importance of each set at the regional scale is reflected by the w_i of the model calibrated with a given set of predictors, while the local scale predictor importance is an output of *biomod2.*

The best performance for the regional study area was obtained with the climatic model (M1, $w_i = 0.820$, Table 2). The second best model, as supported by the occurrence information, was related with landscape composition (M2, $w_i = 0.174$). The null model presented the lowest values of w_i (M5, $w_i = 1.91 \times 10^{-6}$).

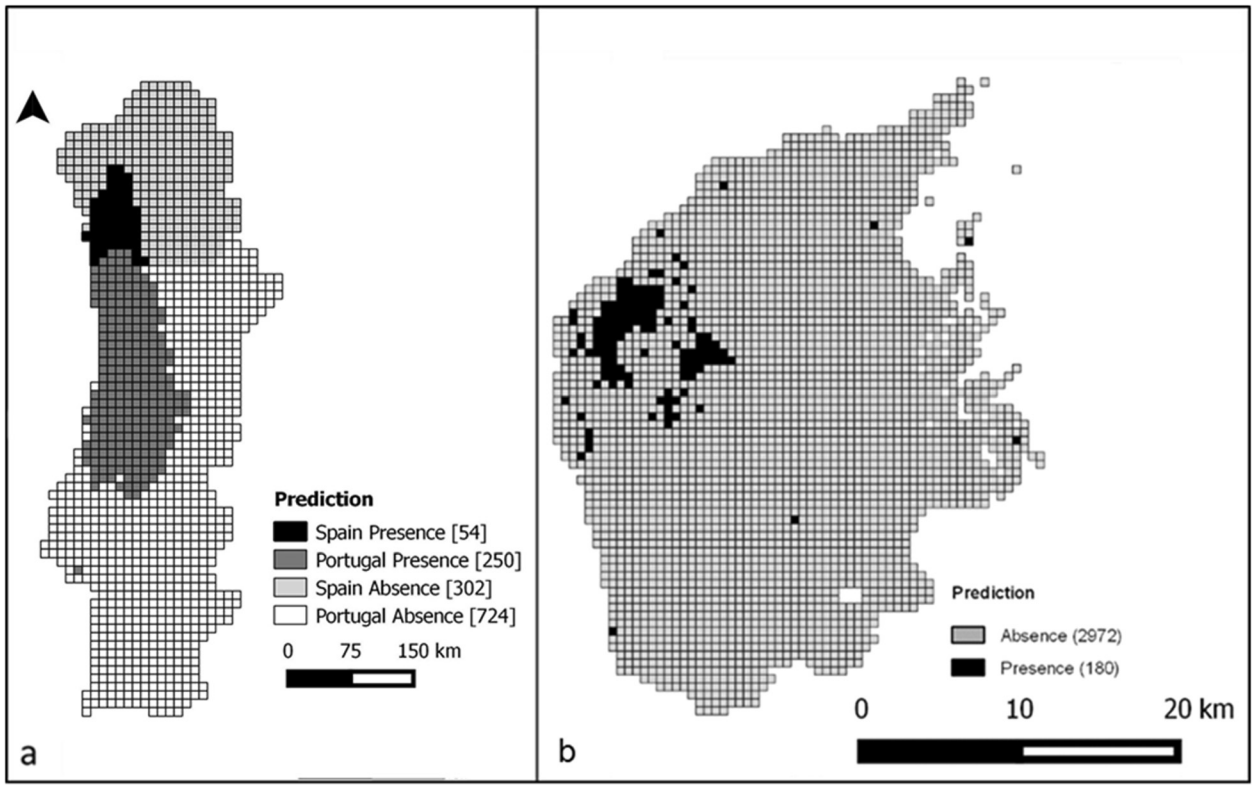

Fig. 3 Potential distribution of *Hakea sericea* in the regional-scale study area obtained by binarization of the average model projection (**a**), and in the local-scale study area, obtained as an ensemble forecast in *biomod2* (**b**)

Discussion

Overcoming scarce occurrence data in invasive species distribution modelling

Scarcity of occurrence data is often a constraint for the calibration of Species Distribution Models, reducing the number of predictors that may be used with confidence (Guisan and Thuiller 2005). This can represent an important limitation to the study and management of biological invasions (Crall et al. 2010). Citizen science programs offer a way of overcoming this problem, by having the interested public participating in data collection. However, while these programs are useful, the resulting data often comprise presence records only, thus yielding unbalanced datasets (Crall et al. 2010).

While ideally situations of low data availability would be addressed by increased sampling, this is not always possible due to economic, temporal or other constraints, and the shortage of high-quality species occurrence data requires the development of improved modelling frameworks, with targeted modifications to deal with specific problems (Lomba et al. 2010; Vicente et al. 2011). For example, the lack of confirmed absence records makes it necessary to use randomly selected pseudo-absences to fit Generalized Linear Models, which require both presence and absence information. To overcome any possible bias created by the random selection of pseudo-absences, the modelling framework was improved through the application of a bootstrapping procedure with a large number of random iterations (5000). Moreover, multimodel inference helped to mitigate the limitation on the number of predictors used to fit the models. Multimodel Inference also provided a way of ranking the importance of predictor sets to explain the distribution of the test species and to generate a robust spatial prediction of regional potential distribution (Vicente et al. 2010).

While useful in situations of deficient species occurrence datasets, the use of a single modelling technique is not ideal; it is well known that different modelling techniques can yield very different results, even when using the same data (Pearson et al. 2006). In cases where more occurrence data is available, as for our local-scale study area, ensemble forecasting may be safely applied (Araújo and New 2007). Applying ensemble modelling allowed us not only to project the potential distribution of *H. sericea* based on a consensus approach, but also to rank the relative importance of the individual environmental predictors, instead of a rank by predictor types obtained with multimodel inference. In our local area, the ranking obtained with *biomod2* identified the presence of foliated

Table 2 Results of Multimodel Inference (MMI) for the regional study area, and importance of each predictor calculated with *biomod2* for the local-scale study area

	Regional scale	Local scale	
	Regional MMI	*biomod2* ensemble	
	W_i	Predictor	Importance
M1 – Climate	0.820	TempRan	0.138
		AnnPrec	0.01
M2 – Landscape composition	0.174	pUrbanA	0.009
		pArtFor	0.015
		pAgrico	0.074
M3 – Landscape structure	0.005	shDiInd	0.059
		NumPatc	0.033
		mShaInd	0.021
		mPeAreR	0.01
M4 – Lithology	0.001	MetRock	0.613
		DetSEdD	0.014
		lithSDI	0.01
M5 – Null model	1.91E-06		
Fire		pMaxBurn	0.067

The values of w_i (always sum up to 1) indicate the likelihood that the model is the best, given the full model and data sets, allowing for a comparison of the importance of each model in explaining the observed distribution of the species. For *biomod2*, the relative importance was calculated for each predictor, indicating its importance in explaining the distribution of *Hakea sericea* in the study area (for more information see Appendix III)
The predictor with the highest importance explaining the local-scale distribution of *H. sericea* was the percentage of foliated metamorphic rocks, or schists (MetRock = 0.613; Table 2). This lithological predictor was followed in importance by temperature annual range (TempRan = 0.138), percentage of agriculture cover (pAgrico = 0.074), Shannon Diversity Index of land cover classes (shDiInd = 0.059), and maximum burnt area per cell (pMaxBurn = 0.067)

metamorphic rocks as the most important predictor explaining the current distribution of *H. sericea*. This was followed in importance by a climatic predictor, confirming the importance of climate conditions in determining the distribution of species (Pearson et al. 2002), even in relatively small areas but with very heterogeneous climate conditions (Vicente et al. 2010).

Measurements of the model's predictive power indicated high accuracy in both the multimodel inference and the ensemble forecasting outputs. By combining the strengths offered by these different modelling approaches, and taking spatial scale into consideration, our framework provides the means of overcoming common difficulties related to data quality and modelling techniques. Such problems include: the risk of over-fitting due to lack of occurrence records (Guisan and Thuiller 2005; addressed here by using Multimodel Inference); uncertainty in model outputs from different techniques (Pearson et al. 2006; addressed here by applying ensemble modelling); variation in the importance of different drivers across spatial scales (Rouget and Richardson 2003; Pearson et al. 2004; Guisan and Thuiller 2005; addressed here by using nested study areas); and difficulty in identifying drivers of distribution masked by stronger gradients (Vicente et al. 2010; dealt with by selecting a local study area based on downscaled predictions of regional distribution models).

As different techniques and variables, better suited for each scale, were used in each analysis, it is important to note that the local-scale analysis cannot be considered a "scaling" of the regional scale, and vice-versa, and it is better interpreted as two different but linked and complementary analyses. With this caveat, our modelling framework enabled us to identify the main sets of drivers of invasion by an aggressive plant species at a coarse spatial scale, and then to rank the importance of individual predictors at a finer scale. It also provided robust spatial predictions of potential distribution for the species at both scales. Considering that prevention is the most cost-effective approach for managing invasive species (Davies and Sheley 2007), the spatial projections obtained in this study provide the means for guiding prevention efforts in environmentally suitable but not yet invaded areas (e.g. large areas of Galicia), thereby providing guidance to efforts directed at surveillance, rapid response and mitigation that are needed to manage transboundary invasions (Hulme 2015). The multiscale nature of our framework means that it may be applied in a workflow where major environmental effects are first identified at a regional scale, and then smaller areas of particular concern (i.e. local scale) are modelled to identify detailed areas and predictors, thus providing insights to inform more directed control efforts (Vicente et al. 2010).

Implications for managing plant invasion in a transboundary context – *Hakea sericea* in the Iberian Peninsula as an illustration

Invasive species management, including prevention and/or control plans, can only be effective if the entire potential distribution of the species in the invaded region is considered (Wilson et al. 2007). In many cases this demands a transboundary approach, when the potential distribution of an invasive species covers two or more neighbouring countries or federal regions, a situation common in many ecological settings (Dallimer and Strange 2015). With this transboundary focus, our work responds to recent calls for the consideration of unaided invasion pathways in the study of invasion (Hulme 2015). This is illustrated in our study by the potential distribution of *H. sericea*, which includes Portugal, where the species is already widespread and considered an aggressive invader, but also Spain, where it is only considered a species with invasive potential. While, to our knowledge, there is only a single reported occurrence of the species in Spain (Pulgar Sañudo 2006), the

spatial projection of our models indicates a considerably larger potential distribution, highlighting the need for investing in surveillance as well as in collaborative prevention and management between the two countries. This is in agreement with the demands of recent European legislation on invasive species (European Parliament and Council of the European Union 2014), demonstrating how modelling frameworks such as ours can offer valuable input for policy and management decisions.

The results obtained regarding the invasion by *H. sericea* in our study area demonstrate the kind of insights a modelling approach such as ours can offer for the transboundary management of an invasive species. Alien species with their native range in areas with mild climates, as is the case of *H. sericea* (Barker et al. 1999), are often limited in newly invaded areas by their frost sensitivity, and for that reason are absent at high elevations (Vicente et al. 2010). Under climate change scenarios, this constraint is expected to be relaxed in the future (Walther 2002) which means that these areas may become more susceptible to invasion. At the same time, the importance of landscape composition predictors, at both regional and local scales, suggests that future shifts in land use could further drive expansion (or contraction) of *H. sericea*'s range in mountainous areas, where some of the most important protected areas in the region occur.

Previous field observations by the authors in Portugal suggest that the species has a preference for areas with schistose bedrock, an observation corroborated by the results presented here. The ability to produce proteoid roots means that *H. sericea* is well adapted to phosphorus-poor soils (Sousa et al. 2007). This may have contributed to its ability to invade by outcompeting native vegetation in areas where phosphorus availability is a key limiting factor, such as those derived from some schists, arenites and large floodplain deposits (Salminen et al. 2005). Felsic rocks in northern Portugal, such as most of the granites, are rich in phosphorus minerals (Neiva et al. 2000). Soils that evolved from these rocks are often juvenile soils or, in some cases, they are still in an early stage of development, mainly in higher areas, where weathering mantles are common. The phosphorus minerals from this rocks, mainly apatite, under wet conditions and under low pH values (typical in granite weathering mantles), release phosphorus by slowly weathering of several phosphorus minerals (such as fluoroapatite), in a well know geochemical alteration process (Bernasconi et al. 2011). Therefore, it is expected that these areas have higher concentrations of phosphorous than lower areas dominated by schistose bedrocks, where soils are often thicker, older and more rich in organic matter. According to the theory of soil development, total soil P and available amounts of mineral P tend to decrease with time (Walker and Syers 1976; Wardle et al. 2004; Menge et al. 2012). For this reason, areas around granitic rocks tend to be less vulnerable to *H. sericea* invasion. The fact that several endemic plants of conservation concern, such as the Dipsacaceae *Succisa pinnatifida* Lange, have a similar preference for areas with schistose bedrock and open scrub vegetation further highlights the need for effective management of *H. sericea*.

While the ranking of predictor importance for the local scale area did not indicate fire as being one of the most important predictors, previous studies have shown an important relationship between *Hakea sericea* and fire, as this species has been observed to cause changes in fire-related variables in some environments (van Wilgen and Richardson 1985), and is at the same time a fire-adapted serotinous species (Groom and Lamont 1997; Brown and Whelan 1999) that releases a large number of seeds after fires. These are mainly dispersed over short distances (Le Maitre et al. 2008), suggesting that fire may be more important in driving local invasion dynamics in the invasive range. The current local invasion of *H. sericea* in Spain is thought to have been triggered by a major fire (Pulgar Sañudo 2006). Exposure to intense natural fires (following a period of fire exclusion) was shown to trigger the invasion of an ecologically-similar serotinous shrub in the Proteaceae family (*Banksia ericifolia*) in South African fynbos (Geerts et al. 2013). The low importance of fire in our models is therefore puzzling, and may be attributable to the particular variable that was used in the model to indicate fire (see discussion in the section on "Predictor selection" above). More work is needed to elaborate the role of fire as a driver of *H. sericea* invasions at different spatial scales in the Iberian Peninsula.

Conclusions

By identifying the potential distribution and its main determinant factors for a given invader, even based on scarce occurrence data, the framework presented here provides the foundation for prioritizing the early management of invasions over large regions. In fact, it allows using the data from a heavily invaded country to predict potential risk areas in a neighbouring country with little or no invasion, provided that model transferability principles are observed (Elith and Leathwick 2009). It is then possible to use the second step of the framework to zoom in areas of particular risk or ecological interest. This may be combined with additional knowledge about the biology of the target species to direct specific management interventions. For example, for *H. sericea* special attention should be given to wildfire occurrence and post-fire invasion dynamics in schistose areas, with management targeted at preventing ecological regime shifts (Gaertner et al. 2014).

Appendix 1

Table 3 Predictors classified into each environmental set, their format and sources for each scale

Class	Predictor	Regional scale		Local scale	
		Format	Source	Format	Source
Climate	MinTemp	Raster (1 km^2 pixel)	http://www.worldclim.org	Raster (1 km^2 pixel)	http://www.worldclim.org
	TempRan				
	AnnPrec				
	PrecSea				
Landscape composition	pUrbanA	Raster	CORINE Land Cover Map (2006)	Vector (Polygon)	COS 2007
	pAgrico				http://www.dgterritorio.pt/cartografia_e_geodesia/cartografia/cartografia_tematica/carta_de_ocupacao_do_solo__cos_/cos__2007
	pArtFor				
	pShrubs				
Landscape structure	NumPatc	Raster	CORINE Land Cover Map (2006)	Vector (Polygon)	COS 2007
	mShaInd				http://www.dgterritorio.pt/cartografia_e_geodesia/cartografia/cos/cos__2007/
	mPerAreR				
	shDiInd				
Lithology	IgnRock	Raster (2500 m^2 pixel)	OneGeology Europe http://www.onegeology-europe.org/home	Raster (2500 m^2 pixel)	OneGeology Europe http://www.onegeology-europe.org/home
	MetRock				
	DetSedD				
	lithSDI				
Fire	pMaxBurn			Vector (polygon)	ICNF http://www.icnf.pt/portal/florestas/dfci/inc/info-geo

Appendix 2

Table 4 Predictors used in each of the models calibrated for multimodel inference (MMI) to evaluate the relative importance of each set of predictors in determining the distribution of *Hakea sericea*

	MinTemp	TempRan	AnnPrec	PrecSea	pUrbanA	pAgrico	pArtFor	pShrubs	NumPatc	mShaInd	mPeAreR	shDiInd	IgnRock	MetRock	DetSEdD	lithSDI	Random predictor
M1. Climate	■	■	■	■													
M2. Landscape composition					■	■	■	■									
M3. Landscape structure									■	■	■	■					
M4. Lithology													■	■	■	■	
M5. Null Model																	■

Predictors (see detailed information in Table 1 and Annex II) are represented in the columns, and models in the rows. The first four models (M1, M2, M3, and M4) reflect each set of environmental predictors, and the last (M5 Null Model) corresponds to a random model used for comparison purposes

Appendix 3

Table 5 Results of Multimodel Inference for the Regional and Local study areas, detailing number of model parameters (k), Akaike weights (w_i), the AICc differences (Δi), and Nagelkerke's R squared (R^2)

	Regional-scale area				Local-scale area		
	k	w_i	Δ*i*	R^2	w_i	Δ*i*	R^2
M1 – Climate	10	0.820	1.241	0.548	0.278	1.908	0.346
M2 – Landscape composition	10	0.174	14.329	0.444	0.722	0	0.361
M3 – Landscape structure	10	0.005	33.719	0.264	1.04E-07	31.510	0.117
M4 – Lithology	10	0.001	41.544	0.181	5.41E-06	23.603	0.178
M5 – Null model	4	1.91E-06	47.461	0.008	8.25E-09	41.159	0.021

Additional file

Additional file 1: Ensemble modelling. (DOCX 16 kb)

Acknowledgements
This work was funded by FEDER funds through the Operational Programme for Competitiveness Factors - COMPETE and by National Funds through FCT - Foundation for Science and Technology under the project PTDC/AAG-MAA/4539/2012 / FCOMP-01-0124-FEDER-027863 (IND_CHANGE). J. Vicente is supported by POPH/FSE funds and by National Funds through FCT - Foundation for Science and Technology through Post-doctoral grant SFRH/BPD/84044/2012. D.M. Richardson acknowledges support from the DST-NRF Centre of Excellence for Invasion Biology and the National Research Foundation (grant 85417).

Availability of data
Data from the invasoras.pt project is available at http://www.invasoras.uc.pt/mapa-de-avistamentos.

Authors' contributions
JM, JV, JH conceived the study. JM, JV, JH, PA, EM, HM, RH collected the data. JM, JV, RH performed statistical analysis. All authors helped to draft manuscript. All authors read and approved the final manuscript.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]InBIO / CIBIO - Centro de Investigação em Biodiversidade e Recursos Genéticos, Universidade do Porto - Campus Agrário de Vairão, Rua Padre Armando Quintas, n° 7, 4485-661 Vairão, Portugal. [2]Centre for Invasion Biology, Department of Botany and Zoology, Stellenbosch University, Matieland 7602, South Africa. [3]Institute of Earth Sciences (ICT/University of Minho/CCT), Minho, Portugal. [4]Centre for Functional Ecology, Department of Life Sciences, University of Coimbra, 3000-456 Coimbra, Portugal. [5]Departamento de Ambiente, Escola Superior Agrária de Coimbra, Instituto Politécnico de Coimbra, Bencanta, 3045-601 Coimbra, Portugal. [6]Faculdade de Ciências da Universidade do Porto, Edifício FC4 (Biologia), Rua do Campo Alegre, s/n, 4169-007 Porto, Portugal. [7]Environmental Resource Management Department, Westlake Conservation Office, Ou Kaapse Weg, Tokai, 7966 City of Cape Town, South Africa.

Received: 20 March 2016 Accepted: 28 June 2016
Published online: 26 July 2016

References
Araújo MB, New M (2007) Ensemble forecasting of species distributions. Trends Ecol Evol 22:42–47. doi:10.1016/j.tree.2006.09.010

Araújo MB, Thuiller W, Williams PH, Reginster I (2005) Downscaling European species atlas distributions to a finer resolution : implications for conservation planning. Glob Ecol Biogeogr 14:17–30. doi:10.1111/j.1466-822X.2004.00128.x

Barker RM, Haegi L, Barker WR (1999) *Hakea sericea*. Flora of Australia Online, http://www.environment.gov.au/biodiversity/abrs/online-resources/flora. Accessed 29 Aug 2014

Bernasconi SM, Bauder A, Bourdon B, Brunner I, Bünemann E, Christl I, Derungs N, Edwards P, Farinotti D, Frey B, Frossard E, Furrer G, Gierga M, Göransson H, Gülland K, Hagedorn F, Hajdas I, Hindshaw R, Ivy-Ochs S, Jansa J, Jonas T, Kiczka M, Kretzschmar R, Lemarchand E, Luster J, Magnusson J, Mitchell EAD, Venterink HO, Plötze M, Reynolds B et al (2011) Chemical and biological gradients along the Damma Glacier soil chronosequence (Switzerland). Vadose Zone J10:867–883

Blackburn TM, Pyšek P, Bacher S, Carlton JT, Duncan RP, Jarošík V, Wilson JRU, Richardson DM (2011) A proposed unified framework for biological invasions. Trends Ecol Evol 26:333–339. doi:10.1016/j.tree.2011.03.023

Breiman L, Friedman J, Stone CJ, Olshen RA (1984) Classification and regression trees. CRC press, New York

Breiman L (2001) Random forests. Mach Learn 45:5–32

Broennimann O, Petitpierre B, Randin C (2015) Ecospat: Spatial Ecology Miscellaneous Methods. Department of Ecology and Evolution (DEE) & Institute of Earth Surface Dynamics (IDYST), University of Lausanne, Switzerland

Broennimann O, Treier UA, Müller-Schärer H, Thuiller W, Peterson AT, Guisan A (2007) Evidence of climatic niche shift during biological invasion. Ecol Lett 10:701–709. doi:10.1111/j.1461-0248.2007.01060.x

Brown CL, Whelan RJ (1999) Seasonal occurrence of fire and availability of germinable seeds in *Hakea sericea* and *Petrophile sessilis*. J Ecol 87:932–941. doi:10.1046/j.1365-2745.1999.00401.x

Brunel S, Schrader G, Brundu G, Fried G (2010) Emerging invasive alien plants for the Mediterranean Basin. EPPO Bull 40:219–238. doi:10.1111/j.1365-2338.2010.02378.x

Burnham K, Anderson D (2002) Model selection and multimodel inference: A practical information-theoretic approach, 2nd edn. Springer, New York

Busby JR (1991) BIOCLIM-a bioclimate analysis and prediction system. Plant Prot Q 6:8–9

Crall AW, Newman GJ, Jarnevich CS, Stohlgren TJ, Waller DM, Graham J (2010) Improving and integrating data on invasive species collected by citizen scientists. Biol Invasions 12:3419–3428. doi:10.1007/s10530-010-9740-9

Dallimer M, Strange N (2015) Why socio-political borders and boundaries matter in conservation. Trends Ecol Evol 30:132–139. doi:10.1016/j.tree.2014.12.004

Davies KW, Sheley RL (2007) A conceptual framework for preventing the spatial dispersal of invasive plants. Weed Sci 55:178–184. doi:10.1614/WS-06-161

Dufour A, Gadallah F, Wagner HH, Guisan A, Buttler A (2006) Plant species richness and environmental heterogeneity in a mountain landscape: effects of variability and spatial configuration. Ecography 29:573–584. doi:10.1111/j.0906-7590.2006.04605.x

EEA (European Environment Agency) (1999) Transboundary air pollution. Environment in the European Union at the turn of the century - Environmental assessment report No 2. pp 133–154EEA (European Environment Agency) (1999) Transboundary air pollution. Environment in the European Union at the turn of the century - Environmental assessment report No 2. pp 133–154.

EEA (European Environment Agency) (2015) The European environment — state and outlook 2015 — synthesis report. European Environmental Agency, Copenhagen

Elith J, Leathwick JR (2009) Species distribution models: Ecological explanation and prediction across space and time. Annu Rev Ecol Evol Syst 40:677–697. doi:10.1146/annurev.ecolsys.110308.120159

Espírito Santo MD, Arsénio P (1999) O género *Hakea* Schrad. em Portugal. 1° Encontro sobre Invasoras Lenhosas. SPCF/ADERE, Gerês, pp 58–65

ESRI (2014) ArcGIS Desktop: release 10.2. Environmental Systems Research Institute, Redlands

European Parliament and Council of the European Union (2014) Regulation (EU) no 1143/2014 of the European Parliament and of the Council of 22 October 2014 on the Prevention and Management of the Introduction and Spread of Invasive Alien Species

Fernandes RF, Vicente JR, Georges D, Alves P, Thuiller W, Honrado JP (2014) A novel downscaling approach to predict plant invasions and improve local conservation actions. Biol Invasions 16:2577–2590. doi:10.1007/s10530-014-0688-z

Friedman JH (1991) Multivariate adaptive regression splines. Ann Stat 19:1–67

Gaertner M, Biggs R, Te Beest M, Hui C, Molofsky J, Richardson DM (2014) Invasive plants as drivers of regime shifts: identifying high-priority invaders that alter feedback relationships. Divers Distrib 20:733–744. doi:10.1111/ddi.12182

Geerts S, Moodley D, Gaertner M, Le Roux JJ, McGeoch MA, Muofhe C, Richardson DM, Wilson JRU (2013) The absence of fire can cause a lag phase: The invasion dynamics of *Banksia ericifolia* (Proteaceae). Austral Ecol 38:931–941. doi:10.1111/aec.12035

Groom PK, Lamont BB (1997) Fruit-seed relations in *Hakea*: serotinous species invest more dry matter in predispersal seed protection. Aust J Ecol 22:352–355. doi:10.1111/j.1442-9993.1997.tb00682.x

Guisan A, Edwards T Jr, Hastie T (2002) Generalized linear and generalized additive models in studies of species distributions: setting the scene. Ecol Modell 157:89–100. doi:10.1016/S0304-3800(02)00204-1

Guisan A, Thuiller W (2005) Predicting species distribution: offering more than simple habitat models. Ecol Lett 8:993–1009. doi:10.1111/j.1461-0248.2005.00792.x

Hastie TJ, Tibshirani RJ (1990) Generalized additive models. CRC Press, London

Hastie T, Tibshirani R, Buja A (1994) Flexible discriminant analysis by optimal scoring. J Am Stat Assoc 89:1255–1270

Hirzel AH, Le Lay G, Helfer V, Randin C, Guisan A (2006) Evaluating the ability of habitat suitability models to predict species presences. Ecol Modell 199:142–152. doi:10.1016/j.ecolmodel.2006.05.017

Hulme PE (2015) Invasion pathways at a crossroad: policy and research challenges for managing alien species introductions. J Appl Ecol 52:1418–1424. doi:10.1111/1365-2664.12470

invasoras.pt (2014) Plantas Invasoras em Portugal. http://www.invasoras.uc.pt/mapa-de-avistamentos/. Accessed 24/08/2014

Le Maitre DC, Krug RM, Hoffmann JH, Gordon AJ, Mgidi TN (2008) *Hakea sericea*: Development of a model of the impacts of biological control on population dynamics and rates of spread of an invasive species. Ecol Model 212:342–358. doi:10.1016/j.ecolmodel.2007.11.011

Le Maitre DC, Richardson DM, Chapman RA (2004) Alien plant invasions in South Africa: driving forces and the human dimension. S Afr J Sci 100:103–112. doi:10.1007/978-1-4612-1926-2_8

Liu C, Berry PM, Dawson TP, Pearson RG (2005) Selecting thresholds of occurrence in the prediction of species distributions. Ecography 28:385–393. doi:10.1111/j.0906-7590.2005.03957.x

Lomba A, Pellissier L, Randin C, Vicente J, Moreira F, Honrado J, Guisan A (2010) Overcoming the rare species modelling paradox: A novel hierarchical framework applied to an Iberian endemic plant. Biol Conserv 143:2647–2657. doi:10.1016/j.biocon.2010.07.007

Marchante H, Morais M, Freitas H, Marchante E (2014) Guia Prático para a Identificação de Plantas Invasoras em Portugal. Imprensa da Universidade de Coimbra, Coimbra

Menge DN, Hedin LO, Pacala SW (2012) Nitrogen and phosphorus limitation over long-term ecosystem development in terrestrial ecosystems. PLoS One 7: e42045. doi:10.1371/journal.pone.0042045

Ministerio de Agricultura Alimentacion y Medio Ambiente (2011) Real Decreto 1628/2011, de 14 de noviembre, por el que se regula el listado y catálogo español de especies exóticas invasoras. BOLETÍN OFICIAL DEL ESTADO 298 Sec. I. 132711–132735

Ministério do Ambiente (1999) Decreto-Lei n.° 565/99 de 21 de Dezembro. Diário da República - I Série n.° 295 9100–9115

Nagelkerke N (1991) A note on a general definition of the coefficient of determination. Biometrika 78:691–692. doi:10.1093/biomet/78.3.691

Neiva A, Silva M, Antunes I, Ramos J (2000) Phosphate minerals of some granitic rocks associated quartz veins from northern and central Portugal. J Czech Geol Soc 45:3–4

Olden JD, Lawler JJ, Poff NL (2008) Machine learning methods without tears: a primer for ecologists. Q Rev Biol 83:171–193. doi:10.1086/587826

Pauchard A, Shea K (2006) Integrating the study of non-native plant invasions across spatial scales. Biol Invasions 8:399–413. doi:10.1007/s10530-005-6419-8

Pearson R, Dawson T, Berry P, Harrison P (2002) SPECIES: A spatial evaluation of climate impact on the envelope of species. Ecol Modell 154:289–300. doi:10.1016/S0304-3800(02)00056-X

Pearson RG, Dawson TP, Liu C (2004) Modelling species distributions in Britain: a hierarchical integration of climate and land-cover data. Ecography (Cop) 27: 285–298. doi:10.1111/j.0906-7590.2004.03740.x

Pearson RG, Thuiller W, Araújo MB, Martinez-Meyer E, Brotons L, McClean C, Miles L, Segurado P, Dawson TP, Lees DC (2006) Model-based uncertainty in species range prediction. J Biogeogr 33:1704–1711. doi:10.1111/j.1365-2699.2006.01460.x

Peterson AT (2003) Predicting the geography of species' invasions via ecological niche modeling. Q Rev Biol 78:419–433. doi:10.1086/378926

Phillips SJ, Dudík M, Schapire RE (2004) A maximum entropy approach to species distribution modeling. Proceedings of the twenty-first international conference on Machine learning, Canada, p 83

Pimentel D, Zuniga R, Morrison D (2005) Update on the environmental and economic costs associated with alien-invasive species in the United States. Ecol Econ 52:273–288. doi:10.1016/j.ecolecon.2004.10.002

Pino J, Font X, Carbó J, Jové M, Pallarès L (2005) Large-scale correlates of alien plant invasion in Catalonia (NE of Spain). Biol Conserv 122:339–350. doi:10.1016/j.biocon.2004.08.006

Pulgar Sañudo Í (2006) Aportaciones a la flora del sur de Galicia (NO España). Bot Complut 30:113–116

QGIS Development Team (2014) Quantum GIS Geographic Information System. Open Source Geospatial Foundation Project. http://qgis.osgeo.org accessed 13 Mar 2016

R Development Core Team (2014) R: A Language and Environment for Statistical Computing. Foundation for Statistical Computing, Vienna, Austria

Rejmánek M, Richardson D (2013) Trees and shrubs as invasive alien species–2013 update of the global database. Divers Distrib 19:1093–1094. doi:10.1111/ddi.12075

Richardson D, Pyšek P, Carlton JT (2011) A compendium of essential concepts and terminology in invasion ecology. In: Richardson DM (ed) Fifty Years of Invasion Ecology: The Legacy of Charles Elton. Wiley-Blackwell, Oxford, pp 409–420

Richardson D, Pyšek P, Rejmánek M, Barbour MG, Panetta FD, West CJ (2000) Naturalization and invasion of alien plants: concepts and definitions. Divers Distrib 6:93–107. doi:10.1046/j.1472-4642.2000.00083.x

Richardson DM, Wilgen BW, Mitchell DT (1987) Aspects of the reproductive ecology of four Australian *Hakea* species (Proteaceae) in South Africa. Oecologia 71:345–354. doi:10.1007/BF00378706

Ridgeway G (1999) The state of boosting. Comput Sci Stat 31:172–181

Ripley BD (1996) Pattern recognition and neural networks. Cambridge University Press, Cambridge

Rivas-Martínez S, Penas A, Díaz TE (2004) Biogeographic map of Europe. Cartographic Service, University of León, Spain

Robin X, Turck N, Hainard A et al (2011) pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics 12: 77. doi:10.1186/1471-2105-12-77

Roques A, Auger-Rozenberg M-A, Blackburn TM, Garnas J, Pyšek P, Rabitsch W, Richardson DM, Wingfield MJ, Liebhold AM, Duncan RP (2016) Temporal and interspecific variation in rates of spread for insect species invading Europe during the last 200 years. Biol Invasions 18 (in press) doi: 10.1007/s10530-016-1080-y

Rouget M, Richardson DM (2003) Understanding patterns of plant invasion at different spatial scales: quantifying the roles of environment and propagule pressure. In: Child LE, Brock JH, Brundu G et al (eds) Plant invasions: ecological threats and management solutions. Backhuys Publishers, Leiden, pp 3–15

Salminen R, Plant JA, Reeder S (2005) Geochemical atlas of Europe. Part 1, Background information, methodology and maps. Geological survey of Finland, Espoo, p 526

Song I-J, Hong S-K, Kim H-O, Byun B, Gin Y (2005) The pattern of landscape patches and invasion of naturalized plants in developed

areas of urban Seoul. Landsc Urban Plan 70:205–219. doi:10.1016/j.landurbplan.2003.10.018
Sousa MF, Façanha AR, Tavares RM, Lino-Neto T, Gerós H (2007) Phosphate transport by proteoid roots of *Hakea sericea*. Plant Sci 173:550–558. doi:10.1016/j.plantsci.2007.08.006
Theoharides KA, Dukes JS (2007) Plant invasion across space and time: factors affecting nonindigenous species success during four stages of invasion. New Phytol 176:256–273. doi:10.1111/j.1469-8137.2007.02207.x
Thuiller W (2007) Biodiversity: Climate change and the ecologist. Nature 448:550–552. doi:10.1038/448550a
Thuiller W, Georges D, Engler R (2015) biomod2: Ensemble platform for species distribution modeling. http://cran.r-project.org/web/packages/biomod2/biomod2.pdf. Accessed 19 May 2015
Thuiller W, Lafourcade B, Engler R, Araújo M (2009) BIOMOD - a platform for ensemble forecasting of species distributions. Ecography 32:369–373. doi:10.1111/j.1600-0587.2008.05742.x
United Nations (1997) Glossary of Environment Statistics. United Nations, New York
van Wilgen B, Richardson D (1985) The effects of alien shrub invasions on vegetation structure and fire behaviour in South African fynbos shrublands: a simulation study. J Appl Ecol 22:955–966. doi:10.2307/2403243
Vicente J, Alves P, Randin C, Guisan A, Honrado J (2010) What drives invasibility? A multi-model inference test and spatial modelling of alien plant species richness patterns in northern Portugal. Ecography 33:1081–1092. doi:10.1111/j.1600-0587.2010.6380.x
Vicente JR, Gonçalves J, Honrado JP, Randin CF, Pottier J, Broennimann O, Lomba A, Guisan A (2014) A framework for assessing the scale of influence of environmental factors on ecological patterns. Ecol Complex 20:151–156. doi:10.1016/j.ecocom.2014.10.005
Vicente J, Randin CF, Gonçalves J, Metzger MJ, Lomba Â, Honrado J, Guisan A (2011) Where will conflicts between alien and rare species occur after climate and land-use change? A test with a novel combined modelling approach. Biol Invasions 13:1209–1227. doi:10.1007/s10530-011-9952-7
Vilà M, Basnou C, Pyšek P, Josefsson M, Genovesi P, Gollasch S, Nentwig W, Olenin S, Roques A, Roy D, Hulme PE, DAISIE partners (2010) How well do we understand the impacts of alien species on ecosystem services? A pan-European, cross-taxa assessment. Front Ecol Environ 8:135–144. doi:10.1890/080083
Vilà M, Espinar JL, Hejda M, Hulme PE, Jarošík V, Maron JL, Pergl J, Schaffner U, Sun Y, Pyšek P (2011) Ecological impacts of invasive alien plants: a meta-analysis of their effects on species, communities and ecosystems. Ecol Lett 14:702–708. doi:10.1111/j.1461-0248.2011.01628.x
Walther G (2002) Weakening of climatic constraints with global warming and its consequences for evergreen broad-leaved species. Folia Geobot 37:129–139. doi:10.1007/bf02803195
Walker TW, Syers JK (1976) The fate of phosphorus during pedogenesis. Geoderma 15:1–19
Wardle DA, Walker LR, Bardgett RD (2004) Ecosystem properties and forest decline in contrasting long-term chronosequences. Science (Wash DC) 305:509–513
Wilson JRU, Wilson JRU, Richardson DM, Procheş Ş, Amis MA, Henderson L, Thuiller W (2007) Residence time and potential range: crucial considerations in modelling plant invasion. Divers Distrib 13:11–12. doi:10.1111/j.1472-4642.2006.00302.x
Wisz MS, Guisan A (2009) Do pseudo-absence selection strategies influence species distribution models and their predictions? An information-theoretic approach based on simulated data. BMC Ecol 9:8. doi:10.1186/1472-6785-9-8